Scientific and Technical Information Center

SEARCH REQUEST FORM

Requester's Full Name: _DAVID GUZO_   Examiner # : _70677_ Date: _11/30/05_

Art Unit: _1636_   Phone Number: _2-0767_   Serial Number: _10/696282_

Location (Bldg/Room#): _Remsen 2A59_ (Mailbox #): _2C70_ Results Format Preferred (circle): (PAPER) DISK

****

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Date: _____

Search Topic:
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

For Sequence Searches Only *Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Please run a regular + interference sequence search on SEQ ID NO: 1 and on a nucleic acid sequence which ~~encode~~ encodes SEQ ID NO: 13.

1- 4718 NA
pg 13 - 736 AA

RECEIVED NOV 30 2005 CHEM/BIOTECH DIV. (STIC)

Thanks

2na-05p 12/12/05

OM protein - nucleic search using frame+ p2n model.

```
                        GenCore version 5.1.6
                        Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:            December 12, 2005, 07:07:34 ; Search time 8012 Seconds
                                               (without alignments)
                                               5079.206 Million cell updates/sec

Title:             US-10-696-282-13
Perfect Score:     3989
Sequence:          1 MAADGYLPDWLEDNLSEGIR............NNGLYTEPRPIGTRYLTRPL 736

Scoring table:     BLOSUM62
                   Xgapop 10.0 ,  Xgapext    0.5
                   Ygapop 10.0 ,  Ygapext    0.5
                   Fgapop  6.0 ,  Fgapext    7.0
                   Delop   6.0 ,  Delext     7.0

Searched:          79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters:    158295336

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn21/USPTO_spool_p/US10696282/runat_06122005_133343_8453/app_query.fasta_1.90
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HRAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10696282 @CCN_1_1 11077 @runat_06122005_133343_8453
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=9 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :   Pending_Patents_NA_Main:*

```
 1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:*
 2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:*
 3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:*
 4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
 5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:*
 6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:*
 7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:*
 8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:*
 9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3989	100.0	2211	54	US-10-696-261-12	Sequence 12, Appl
2	3989	100.0	2211	54	US-10-696-282-12	Sequence 12, Appl
3	3989	100.0	2211	54	US-10-696-900-12	Sequence 12, Appl
4	3989	100.0	4718	1	PCT-US02-33629-6	Sequence 6, Appli
5	3989	100.0	4718	1	PCT-US02-38423-20	Sequence 20, Appli
6	3989	100.0	4718	1	PCT-US03-11191-3	Sequence 3, Appli
7	3989	100.0	4718	43	US-10-291-583-6	Sequence 6, Appli
8	3989	100.0	4718	51	US-10-427-129-1	Sequence 1, Appli

OM protein - nucleic search, using frame+ applus p2n model

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

Run on: December 12, 2005, 11:30:19 ; Search time 239 Seconds
(without alignments)
1151.285 Million cell updates/sec

Title: US-10-696-282-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSEGIR.........NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62

Xgapop 10.0 ,	Xgapext 0.5		
Ygapop 10.0 ,	Ygapext 0.5		
Fgapop 6.0 ,	Fgapext 7.0		
Delop 6.0 ,	Delext 7.0		

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 67844860

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US10696282/runat_06122005_133344_8538/app_query.fasta_1.90
-O=/cgn2_1/USPTO_spool_P/US10696282/runat_06122005_133344_8538/app_query.fasta -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10696282 @CGN_1 183 @runat_06122005_133344_8538
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=1 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3396.5	85.1	4679	7	US-11-184-380-25	Sequence 25, Appl
2	2486.5	62.3	2208	6	US-10-719-311-1	Sequence 5, Appl
3	2481.5	62.2	4768	7	US-10-719-311-1	Sequence 19, Appl
4	2220	55.7	4652	7	US-11-184-380-7	Sequence 7, Appl
5	2220	55.7	2307	7	US-11-184-380-1	Sequence 1, Appl
6	2214	55.5	2264	7	US-11-184-380-8	Sequence 8, Appl
7	2214	55.5	2264	7	US-11-184-380-9	Sequence 9, Appl
8	1830.5	45.9	1800	6	US-10-719-311-17	Sequence 17, Appl
9	1690.5	42.4	1617	6	US-10-719-311-19	Sequence 19, Appl
10	136	3.4	3289	7	US-10-933-025-14	Sequence 14, Appl
11	136	3.4	4286	6	US-10-933-025-13	Sequence 13, Appl
12	132	3.3	3906	6	US-10-933-025-19	Sequence 19, Appl
13	127.5	3.2	2616	6	US-10-933-025-5	Sequence 5, Appl
14	127.5	3.2	3363	7	US-10-933-025-4	Sequence 4, Appl
15	119	3.0	191884	7	US-11-121-086-2	Sequence 2, Appl
16	118.5	3.0	5296	6	US-11-121-086-217	Sequence 217, Appl
17	115.5	2.9	3461	6	US-10-131-826A-199	Sequence 199, App
18	112.5	2.8	2245	6	US-10-401-386B-61	Sequence 61, Appl
19	111	2.8	2901	6	US-10-485-517-38	Sequence 38, Appl
20	111	2.8	2901	6	US-10-485-517-54	Sequence 54, Appl
21	111	2.8	169495	7	US-11-121-086-61	Sequence 61, Appl
22	108	2.7	153376	7	US-11-186-284-29	Sequence 29, Appl
23	108	2.7	172543	7	US-11-135-855-21	Sequence 21, Appl
24	107.5	2.7	1068	7	US-10-851-667A-24	Sequence 24, Appl
25	107.5	2.7	5665	6	US-10-821-234-244	Sequence 244, App
26	107	2.7	4749	6	US-10-821-234-244	Sequence 244, App
27	107	2.7	4770	7	US-11-000-463-7	Sequence 7, Appl
28	107	2.7	5921	7	US-11-186-284-29	Sequence 29, Appl
29	107	2.7	6728	7	US-11-186-284-27	Sequence 27, Appl
30	106	2.7	4965	6	US-10-485-517-43	Sequence 43, Appl
31	106	2.7	6763	6	US-11-080-991-55	Sequence 55, Appl
32	105.5	2.6	1155	6	US-11-080-991-55	Sequence 55, Appl
33	105.5	2.6	168516	6	US-10-833-270-1	Sequence 1, Appl
34	104.5	2.6	2436	7	US-10-102-240-99	Sequence 99, Appl
35	104	2.6	55763	6	US-11-102-766-1	Sequence 1, Appl
36	103.5	2.6	2197	7	US-11-000-463-225	Sequence 225, App
37	103	2.6	169495	7	US-11-121-086-61	Sequence 61, Appl
38	103	2.6	5891	6	US-11-186-284-27	Sequence 27, Appl
39	103	2.6	2745	7	US-10-750-185-56614	Sequence 56614, A
40	102.5	2.6	9454	7	US-11-109-056-1	Sequence 1, Appl
41	102	2.6	2349	6	US-10-972-053-7	Sequence 7, Appl
42	102	2.6	2557	6	US-10-972-053-11	Sequence 11, Appl
43	101.5	2.5	2133	7	US-11-045-802-10	Sequence 10, Appl
44	101.5	2.5	3469	6	US-10-933-025-7	Sequence 7, Appl
45	101.5	2.5	7989	6	US-10-509-921-8	Sequence 8, Appl

RESULT 1

US-11-184-380-25
Sequence 25, Application US/11184380

; Publication No. US20050255089A1
; GENERAL INFORMATION:
; APPLICANT: Chiorini, John
; APPLICANT: Kotin, Robert M.
; TITLE OF INVENTION: AAV5 NUCLEIC ACIDS
; FILE REFERENCE: 14014_032303
; CURRENT APPLICATION NUMBER: US/11/184,380
; CURRENT FILING DATE: 2005-07-19
; PRIOR APPLICATION NUMBER: PCT/US99/11958
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,029
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 4679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Note =
US-11-184-380-25
; OTHER INFORMATION: synthetic construct

ALIGNMENTS

US-11-184-380-25

Alignment Scores:
Pred. No.:
Score: 3396.50
Pred.: 3.15e-280
Percent Similarity: 90.08%
Percent Identity: 83.15%
Best Local Similarity: 85.15%

Length:	4679		
Matches:	612		
Conservative:	51		
Mismatches:	72		
Indels:	1		

Query Match: 85.15%

OM protein - nucleic search, using frame+p2n model

```
Copyright (c) 1993 - 2005 Compugen Ltd.

                                        GenCore version 5.1.6

Run on:             December 12, 2005, 09:47:27 ; Search time 1262 seconds
                                                   (without alignments)
                                        4822.712 Million cell updates/sec

Title:              US-10-696-282-13
Perfect score:      3989
Sequence:           1 MAADGYLPDWLEDNLSEGIR.........NNGLYTEPRPIGTRYLTRPL 736

Scoring table:      BLOSUM62
    Xgapop  10.0 ,  Xgapext  0.5
    Ygapop  10.0 ,  Ygapext  0.5
    Fgapop   6.0 ,  Fgapext  7.0
    Delop    6.0 ,  Delext   7.0

Searched:           9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters:    19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Database :   Published_Applications_NA_Main:*
-MODEL=frame+_p2n.model
-Q=/cgn2_1/USPTO_spool_p/US10696282/runat_06122005_133344_8507/app_query.fasta_1.90
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-DEV=xlp -UNITS=bits -START=1 -END=-1
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-MATRIX=blosum62 -TRANS=human40.cdi -MODE=LOCAL -OUTFMT=pto -NORM=ext
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -DOCALIGN=200 -THR_SCORE=pct
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10696282 @CGN_1_1549 @runat_06122005_133344_8507 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3989	100.0	2211	7	US-10-696-261-12	Sequence 12, Appl
2	3989	100.0	2211	7	US-10-696-282-12	Sequence 12, Appl
3	3989	100.0	2211	7	US-10-696-900-12	Sequence 12, Appl
4	3989	100.0	4718	6	US-10-291-583-6	Sequence 6, Appli
5	3989	100.0	4718	7	US-10-291-583-1	Sequence 1, Appli
6	3989	100.0	4718	7	US-10-696-282-1	Sequence 1, Appl
7	3989	100.0	4718	7	US-10-696-900-1	Sequence 1, Appl

```
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

ALIGNMENTS

RESULT 1
US-10-696-261-12

```
; Sequence 12, Application US/10696261
; Publication No. US2004005731A1
; GENERAL INFORMATION:
;  APPLICANT: Wilson, James M.
;  APPLICANT: Xiao, Weidong
;  TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
;  TITLE OF INVENTION: Vectors and Host Cells Containing Same
;  FILE REFERENCE: GNVPN.031USA
;  CURRENT APPLICATION NUMBER: US/10/696,261
;  CURRENT FILING DATE: 2003-10-29
;  PRIOR APPLICATION NUMBER: US/09/807,802A
;  PRIOR FILING DATE: 2002-02-21
;  PRIOR APPLICATION NUMBER: US 60/107,114
;  PRIOR FILING DATE: 1998-11-05
;  PRIOR APPLICATION NUMBER: PCT/US99/25694
;  PRIOR FILING DATE: 1999-11-02
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;  LENGTH: 2211
;  TYPE: DNA
;  ORGANISM: AAV-1
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(2208)
;  OTHER INFORMATION:
US-10-696-261-12
```

Alignment Scores:

```
  8   3989   100.0   4718   7   US-10-427-129-1   Sequence 1, Appli
  9   3989   100.0   4718   7   US-10-959-017-3   Sequence 3, Appli
 10   3989   100.0  13804   7   US-10-427-122-11  Sequence 11, Appl
 11   3984   99.9    4718   7   US-10-216-870-11  Sequence 11, Appli
 12   3984   99.9    7447   8   US-10-415-834-11  Sequence 11, Appl
 13   3963   99.3    4683   7   US-10-427-122-6   Sequence 6, Appli
 14   3963   99.3    4683   9   US-10-959-017-2   Sequence 2, Appli
 15   3945   98.9    4683   9   US-10-696-261-19  Sequence 19, Appli
 16   3945   98.9    4683   9   US-10-696-282-19  Sequence 19, Appl
 17   3945   98.9    4683   7   US-10-696-900-19  Sequence 19, Appl
 18   3511   88.0    4722   7   US-10-427-129-3   Sequence 3, Appli
 19   3494   87.6    4726   9   US-10-496-799-1   Sequence 1, Appli
 20   3494   87.6    4726   9   US-10-291-583-8   Sequence 8, Appli
 21   3473   87.1    3122   8   US-10-291-583-45  Sequence 45, Appl
 22   3467   86.9    3122   6   US-10-291-583-45  Sequence 45, Appl
 23   3464   86.8    4382   8   US-10-291-583-43  Sequence 43, Appli
 24   3460   86.8    3117   6   US-10-291-583-44  Sequence 44, Appli
 25   3457   86.7    3121   6   US-10-291-583-42  Sequence 42, Appli
 26   3453   86.6    3122   6   US-10-291-583-27  Sequence 27, Appli
 27   3451.5 86.5    3128   6   US-10-291-583-26  Sequence 26, Appli
 28   3448   86.4    3075   6   US-10-291-583-59  Sequence 59, Appl
 29   3446.5 86.4    3129   6   US-10-291-583-25  Sequence 25, Appli
 30   3445   86.4    3123   6   US-10-291-583-41  Sequence 41, Appli
 31   3442.5 86.3    3142   6   US-10-291-583-11  Sequence 11, Appli
 32   3441   86.3    4721   6   US-10-291-583-1   Sequence 1, Appli
 33   3440   86.3    3123   6   US-10-291-583-47  Sequence 47, Appli
 34   3439   86.2    3128   6   US-10-291-583-28  Sequence 28, Appli
 35   3438   86.2    8179   5   US-10-291-942-5   Sequence 5, Appli
 36   3434   86.2    3084   6   US-10-291-583-39  Sequence 39, Appli
 37   3432   86.1    3098   6   US-10-291-583-37  Sequence 37, Appli
 38   3429   86.0    3197   6   US-10-291-583-29  Sequence 29, Appli
 39   3424.5 85.8    3142   6   US-10-291-583-11  Sequence 11, Appli
 40   3423.5 85.8    3121   6   US-10-291-583-15  Sequence 15, Appli
 41   3423.5 85.8    2205   6   US-10-291-583-120 Sequence 120, App
 42   3422.5 85.8    4385   6   US-10-291-583-5   Sequence 5, Appli
 43   3422   85.8    3095   6   US-10-291-583-24  Sequence 24, Appli
 44   3420.5 85.7    3121   6   US-10-291-583-13  Sequence 13, Appli
 45   3420   85.7    3113   6   US-10-291-583-31  Sequence 31, Appli
         3098   6   US-10-291-583-9   Sequence 9, Appli
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using "frame+plus_p2n model"
(without alignments)

Run on: December 12, 2005, 06:54:48 ; Search time 331 seconds
3952.524 Million cell updates/sec

```
Title:          US-10-696-282-13
Perfect score:  3989
Sequence:       1 MAADGYLPDWLEDNLSEGIR.........NNGLYTEPRPIGTRYLTRPL 736

Scoring table:  BLOSUM62
                Xgapop 10.0 ,   Xgapext   0.5
                Ygapop 10.0 ,   Ygapext   0.5
                Fgapop  6.0 ,   Fgapext   7.0
                Delop   6.0 ,   Delext    7.0

Searched:       1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters:   2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10696282/runat_06122005_133343_8433/app_query.fasta_1.90
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10696282 @CGN_1_1_290 @runat_06122005_133343_8433 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

      1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
      2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
      3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
      4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
      5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
      6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
      7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
      8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
      9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query	Match Length	DB	ID	Description
1	3989	100.0	2211	3	US-09-807-802A-12	Sequence 12, Appli
2	3989	100.0	4718	3	US-09-807-802A-1	Sequence 1, Appli
3	3984	99.9	7447	3	US-10-216-870-11	Sequence 11, Appl
4	3945	98.9	4683	3	US-09-807-802A-19	Sequence 19, Appl
5	3439	86.2	8179	3	US-09-438-268-5	Sequence 5, Appli
6	3402.5	85.3	4072	3	US-09-770-315-4	Sequence 4, Appl
7	3402.5	85.3	4679	3	US-10-038-972A-12	Sequence 12, Appl
8	3402.5	85.3	7557	3	US-09-770-315-3	Sequence 3, Appli
9	3402.5	85.3	8698	3	US-09-770-315-2	Sequence 2, Appli
10	3360.5	84.2	4680	2	US-08-254-358-1	Sequence 1, Appli
11	3360.5	84.2	4680	2	US-08-475-391-1	Sequence 1, Appli
12	3360.5	84.2	4680	2	US-08-709-609-1	Sequence 1, Appli
13	3360.5	84.2	4680	6	PCT-US95-07178-1	Sequence 1, Appli
14	3348.5	83.9	4681	3	US-09-807-802A-18	Sequence 18, Appl
15	3322.5	83.3	4675	3	US-09-782-378A-1	Sequence 1, Appli
16	3322.5	83.3	4675	3	US-09-782-378A-2	Sequence 2, Appli
17	3251	81.5	1800	3	US-09-807-802A-14	Sequence 14, Appl
18	3038.5	76.2	8151	3	US-09-438-268-2	Sequence 2, Appli
19	2906	72.9	7744	3	US-10-216-870-14	Sequence 14, Appl
20	2494.5	62.5	1605	3	US-09-807-802A-16	Sequence 16, Appl
21	2486.5	62.3	2208	3	US-09-532-594B-5	Sequence 5, Appli
22	2486.5	62.3	7214	3	US-10-216-870-1	Sequence 1, Appli
23	2481.5	62.2	4767	3	US-09-532-594B-1	Sequence 1, Appli
24	2220	55.7	2307	3	US-09-533-427-7	Sequence 7, Appli
25	2220	55.7	4652	3	US-09-533-427-1	Sequence 1, Appli
26	2214	55.5	2264	3	US-09-533-427-8	Sequence 8, Appli
27	2214	55.5	2264	3	US-09-533-427-9	Sequence 9, Appli
28	1830.5	45.9	1800	3	US-09-532-594B-17	Sequence 17, Appl
29	1690.5	42.4	1617	3	US-09-532-594B-19	Sequence 19, Appl
30	1410	35.3	2271	3	US-09-438-268-3	Sequence 3, Appli
31	599.5	15.0	2380	3	US-09-533-427-26	Sequence 26, Appl
32	599.5	15.0	2380	3	US-10-187-253E-26	Sequence 26, Appl
33	599.5	15.0	4678	3	US-10-187-253E-32	Sequence 32, Appl
34	599.5	15.0	4678	3	US-10-187-253E-22	Sequence 22, Appl
35	486	12.2	1699	3	US-10-187-253E-23	Sequence 23, Appl
36	486	12.2	1699	3	US-10-187-253E-28	Sequence 28, Appl
37	439	11.0	5049	2	US-08-336-345-34	Sequence 34, Appl
38	439	11.0	5049	2	US-08-336-345-1	Sequence 1, Appli
39	439	11.0	5049	2	US-08-647-655-1	Sequence 1, Appli
40	439	11.0	5049	2	US-08-647-655-2	Sequence 2, Appli
41	277.5	7.0	681	3	US-10-187-253E-92	Sequence 92, Appl
42	263	6.6	2062	9	US-10-187-253B-22	Sequence 22, Appl
43	247	6.2	1740	2	US-07-969-213-1	Patent No. 5223424-3
44	235.5	5.9	700	3	US-10-187-253B-4	Sequence 4, Appli
45	233.5	5.9	700	3	US-10-187-253E-5	Sequence 5, Appli

RESULT 1
US-09-807-802A-12
Patent No. 6759237
Sequence 12, Application US/09807802A

GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVPN_031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1

; SEQ ID NO 12
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
; OTHER INFORMATION:
US-09-807-802A-12

Alignment Scores:
Pred. No.:   0
Score: 3989.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%

Length: 2211
Matches: 2211
Conservative: 0
Mismatches: 0

ALIGNMENTS

OM protein - nucleic search, using frame+plus_p2n model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: December 12, 2005, 06:00:53 ; Search time 6192 Seconds
(without alignments)
5561.255 Million cell updates/sec

Title: US-10-696-282-13
Perfect score: 3989
Sequence: 1 MAADGYLPDWLEDNLSEGIR...........NNGLYTEPPIGTRYLTRPL 736

Searched: 41078325 seqs, 2339541228 residues

Total number of hits satisfying chosen parameters: 82156650

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10692282//runat_06122005_133342_8420/app_query.fasta_1.90
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10692282 @CGN_1_8010 @runat_06122005_133342_8420 -NCPU=6 -LONGLOG
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est6:*
7: gb_est7:*
8: gb_gss1:*
9: gb_gss2:*
10: gb_gss3:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	415	10.4	264	7	CO92248	BovGen_20
c 2	294	7.4	195	7	CO88893	BovGen_17
c 3	193.5	4.9	753	9	BH115587	RPCI-24-3
4	165	4.1	1944	10	AY415352	Mus muscu
5	151.5	3.8	572	5	BW334595	BW334595
6	148	3.7	1109	8	DR147327	49053839
7	147	3.7	2318	4	BC014681	Homo sapi
8	139.5	3.5	803	9	BZ265284	CH230-375
9	136.5	3.5	1297	4	CNS0FQWD	
10	136	3.4	3945	4	CR684225	Tetraodon
11	135	3.4	667	3	CR749319	Homo sapi
12	134	3.4	856	9	BW323619	BW323619
13	133	3.3	642	5	BH164736	ENTIS28TF
14	133	3.3	1947	10	BW360495	BW360495
15	133	3.3	2146	4	AY415350	Homo sapi
16	133	3.3	2295	4	AK042727	Mus muscu
17	130.5	3.3	2990	4	AK035953	Mus muscu
18	130	3.2	687	5	CR857451	Pongo pyg
19	129	3.2	3943	4	BW482915	BW482915
20	128.5	3.2	552	6	AK087510	Mus muscu
21	128	3.2	4355	10	CF318715	HD--08-P0
22	127.5	3.2	921	8	AY416870	Mus muscu
23	127.5	3.2	1902	4	CG239378	OG2CN09TV
24	127.5	3.2	2613	11	CR602498	full-leng
25	127.5	3.2	5021	4	DQ052183	Homo sapi
c 26	125	3.2	640	6	CR857584	Pongo pyg
c 27	125	3.1	701	6	CA356072	NC
c 28	125	3.1	703	6	CA363789	638563 NC
c 29	125	3.1	745	3	CA378088	656813 NC
30	125	3.1	754	2	BJ139005	BJ139005
31	125	3.1	2019	4	BG207226	RST26694
32	124	3.1	925	5	AK014322	Mus muscu
33	123.5	3.1	645	5	BQ440075	BQ440075
c 34	123.5	3.1	1329	4	BI911451	603063370
35	123.5	3.1	3512	4	CR591190	full-leng
36	123.5	3.1	4763	4	AK083225	Mus muscu
37	123	3.1	631	5	BW328301	BW328301
38	123	3.1	6987	4	BC069829	Homo sapi
c 39	123	3.1	703	7	CF869635	trIc019xp
c 40	122.5	3.1	756	9	AQ742711	HS 5386 B
c 41	122.5	3.1	795	6	CB899782	trIc019xp
42	122.5	3.1	2391	10	CL961669	OSIPCCC37
43	122.5	3.1	2610	11	DQ052184	Pan trogl
44	122	3.1	4180	4	AL133028	Homo sapi
45	121.5	3.0	925	6	BH152154	ENTPV12FR

RESULT 1
CO892248/c

LOCUS CO892248 264 bp mRNA linear EST 01-SEP-2004
DEFINITION BovGen 20573 normal cattle brain Bos taurus cDNA clone
RZPDp1056M03600 5', mRNA sequence.
ACCESSION CO892248
VERSION CO892248.1 GI:51822548
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 264)
AUTHORS Hennig,S., Janitz,M., Herwig,R. and Williams,J.
TITLE Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
JOURNAL Unpublished (2004)
COMMENT Contact: Hennig S
laboraty 123, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Ihnestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer

ALIGNMENTS

```
                         GenCore version 5.1.6
                         Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, Listing 'frame+plus_p2n.model'
                                              (without alignments)
Run on:          December 12, 2005, 00:05:58 ; Search time 974 Seconds
                                              5036.154 Million cell updates/sec

Title:          US-10-696-282-13
Perfect score:  3989
Sequence:       1 MAADGYLPDWLEDNLSEGIR.........NNGLYTEPRPIGTRYLTRPL 736

Scoring table:  BLOSUM62
                Xgapop   10.0  ,  Xgapext   0.5
                Ygapop   10.0  ,  Ygapext   0.5
                Fgapop    6.0  ,  Fgapext   7.0
                Delop     6.0  ,  Delext    7.0

Searched:       4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters:      9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10696282/runat_06122005_133341_8403/app_query.fasta_1.90
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10696282 @CGN_1_1_1096 @runat_06122005_133341_8403 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases : N_Geneseq_21:*
     1:     geneseqnl980s:*
     2:     geneseqnl990s:*
     3:     geneseqn2000s:*
     4:     geneseqn2001as:*
     5:     geneseqn2001bs:*
     6:     geneseqn2002as:*
     7:     geneseqn2002bs:*
     8:     geneseqn2003as:*
     9:     geneseqn2003bs:*
    10:     geneseqn2003cs:*
    11:     geneseqn2003ds:*
    12:     geneseqn2004as:*
    13:     geneseqn2004bs:*
    14:     geneseqn2005s:*
```

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	3989	100.0	2211	3	AAD00777	Adeno-ass
2	3989	100.0	2211	14	AD227052	Adeno-ass
3	3989	100.0	2211	14	AD226929	Adeno-ass
4	3989	100.0	4347	13	ADW39398	Adeno-ass

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
   5    3989  100.0  4347  14  Adz446594  HSV-AAV 8
   6    3989  100.0  4718   3  Aad00772   Adeno-ass
   7    3989  100.0  4718  10  Ade76507   Adeno-ass
   8    3989  100.0  4718  14  Adl13984   Adeno-ass
   9    3989  100.0  4718  12  Adg39758   AAV-1 gen
  10    3989   99.9  7447   8  Acc58477   Vector pP
  11    3968   99.5  2211  14  Adz226928  Adeno-ass
  12    3963   99.3  2211  14  Adz27053   Adeno-ass
  13    3963   99.3  4239  14  Adw39402   Adeno-ass
  14    3963   99.3  4239  14  Adz446598  HSV-AAV 8
  15    3963   99.3  4683   4  Aaf23749   AAV6 DNA
  16    3963   96.6  4683  10  Adl13983   Adeno-ass
  17    3963   99.3  4683  12  Adg39763   AAV-6 gen
  18    3883.5 97.4  2214  14  Adz226930  Adeno-ass
  19    3873   97.1  2211  14  Adz226931  Adeno-ass
  20    3853   96.6  4722   4  Aaf23748   AAV3B DNA
  21    3511   88.0  4722  12  Adg39760   AAV-3B ge
  22    3511   88.0  4722  14  Adz27050   Adeno-ass
  23    3494   87.6  2211  14  Adz226932  Adeno-ass
  24    3494   87.6  4286  14  Adz446595  HSV-AAV 8
  25    3494   87.6  4287  13  Adw39399   Adeno-ass
  26    3494   87.6  4726  10  Ade76509   Adeno-ass
  27    3494   87.6  4726  12  Adl13985   Adeno-ass
  28    3494   87.6  4726  14  Adg39759   AAV-3A ge
  29    3483.5 87.3  2214  14  Adz226891  Adeno-ass
  30    3473   87.1  3122  14  Ade76546   Adeno-ass
  31    3469   87.0  2217  14  Adz226892  Adeno-ass
  32    3467   86.9  4382   9  Abq80410   AAV9 rep
  33    3464   86.8  3117  10  Ade76544   Adeno-ass
  34    3460   86.7  3121  10  Ade76545   Adeno-ass
  35    3458.5 86.7  2210  14  Adz226886  Adeno-ass
  36    3457   86.7  3122  14  Ade76543   Adeno-ass
  37    3456.5 86.7  2214  14  Adz226885  Adeno-ass
  38    3456.5 86.7  2214  14  Adz226890  Adeno-ass
  39    3456.5 86.7  3158  14  Adz27020   Adeno-ass
  40    3456   86.6  2217  14  Adz226869  Adeno-ass
  41    3455.5 86.6  2214  14  Adz226882  Adeno-ass
  42    3453.5 86.6  2208  14  Adz226895  Adeno-ass
  43    3453.5 86.6  2208  14  Adz226899  Adeno-ass
  44    3453   86.6  2217  14  Adz226863  Adeno-ass
  45    3453   86.6  2217  14  Adz226859  Adeno-ass
```

ALIGNMENTS

```
RESULT 1
AAD00777
ID   AAD00777 standard; DNA; 2211 BP.
XX
AC   AAD00777;
XX
DT   08-SEP-2000  (first entry)
XX
DE   Adeno-associated virus serotype 1 capsid protein VP1 DNA.
XX
KW   Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW   cap protein; recombinant viral vector; gene delivery; gene therapy;
KW   vaccine; transgene; VP1; ss.
XX
OS   Adeno associated virus serotype 1.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..2211
FT                   /tag= a
FT                   /product= "VP1 protein"
XX
PN   WO200028061-A2.
XX
PD   18-MAY-2000.
XX
PF   02-NOV-1999;  99WO-US025694.
XX
PR   05-NOV-1998;  98US-0107114P.
```

```
OM protein - nucleic search, using frameplus p2n model
                                     without alignments

                    GenCore version 5.1.6
              Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:          December 12, 2005, 05:57:53 ; Search time 8511 Seconds
                                                (without alignments)
                                    4915.613 Million cell updates/sec

Title:           US-10-696-282-13
Perfect score:   3989
Sequence:        1 MAADGYLPDWLEDNLSEGIR...........NNGLYTEPRPIGTRYLTRPL 736

Scoring table:   BLOSUM62
                 Xgapop 10.0 ,  Xgapext  0.5
                 Ygapop 10.0 ,  Ygapext  0.5
                 Fgapop  6.0 ,  Fgapext  7.0
                 Delop   6.0 ,  Delext   7.0

Searched:        5883141 seqs,  28421725653 residues

Total number of hits satisfying chosen parameters:   11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10696282/runat_06122005_133342_8409/app_query.fasta_1.90
-DB=GenEmbl -QFMT=fastap -SUFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10696282 @CGN 1 1 7415 @runat 06122005 133342 8409 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -YGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       GenEmbl:.*
                 1:  gb_ba:*
                 2:  gb_in:*
                 3:  gb_env:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pr:*
                 9:  gb_ro:*
                10:  gb_sts:*
                11:  gb_sy:*
                12:  gb_un:*
                13:  gb_vi:*
                14:  gb_htg:*
                15:  gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                           SUMMARIES

Result               Query
 No.     Score  Match Length DB ID         Description
------   -----  ----- ------ -----         -----------
  1      3989   100.0  2211   6  BD242771   Adeno-ass
  2      3989   100.0  2211   6  CS073491   Sequence
  3      3989   100.0  2211   6  CS073614   Sequence
```

```
   4   3989 100.0  2211   6  AR562503  Sequence
   5   3989 100.0  4718   6  BD242766  Adeno-ass
   6   3989 100.0  4718   6  AR562498  Sequence
   7   3989 100.0  4718   6  AX753251  Sequence
   8   3989 100.0  4718  13  AF063497  Adeno-ass
   9   3984  99.5  7447   6  AR527492  Sequence
  10   3968  99.5  2211   6  CS073490  Sequence
  11   3968  99.5  2211  13  AY530611  Adeno-ass
  12   3963  99.3  2211   6  CS073615  Sequence
  13   3963  99.3  4683  13  AF028704  Adeno-ass
  14   3945  98.9  4683   6  BD242775  Adeno-ass
  15   3945  98.9  4683   6  AR562507  Sequence
  16   3883.5 97.4 4718   6  CS073492  Sequence
  17   3883.5 97.4 2214  13  AY530606  Adeno-ass
  18   3873  97.1  2211   6  CS073493  Sequence
  19   3873  97.1  2211  13  AY530607  Adeno-ass
  20   3853  96.6  2211   6  CS073494  Sequence
  21   3853  96.6  2211  13  AY530609  Adeno-ass
  22   3511  88.0  4722  13  AF028705  Adeno-ass
  23   3494  87.6  2211   6  CS073612  Sequence
  24   3494  87.6  4726  13  AX753253  Sequence
  25   3494  87.6  4726  13  AV748704  Adeno-assoc
  26   3483.5 87.3 2214   6  CS073453  Sequence
  27   3483.5 87.3 2214  13  AX242997  Non-human
  28   3473  87.1  3122  13  AX753290  Sequence
  29   3469.5 87.1 4430  13  AY695376  Adeno-ass
  30   3469  87.0  3122  13  AX753290  Sequence
  31   3469  87.0  2217   6  CS073454  Sequence
  32   3464  86.8  3117   6  AX753288  Sequence
  33   3460  86.7  3121   6  AX753289  Sequence
  34   3458.5 86.7 3121   6  AX753448  Sequence
  35   3457  86.7  3122  13  AX753287  Sequence
  36   3456.5 86.7 2208  13  AY530622  Adeno-ass
  37   3456.5 86.7 2208   6  CS073447  Sequence
  38   3456.5 86.7 2214   6  CS073452  Sequence
  39   3456.5 86.7 2214  13  AY530567  Adeno-ass
  40   3456   86.7 3158   6  CS073582  Sequence
  41   3456   86.6 2217   6  CS073431  Sequence
  42   3455.5 86.6 2214   6  CS073444  Sequence
  43   3455.5 86.6 2214  13  AY530561  Adeno-ass
  44   3453.5 86.6 2208   6  CS073461  Sequence
  45   3453.5 86.6 2208  13  AY530591  Adeno-ass
```

                              ALIGNMENTS

```
RESULT 1
BD242771
LOCUS       BD242771           2211 bp   DNA    linear   PAT 17-JUL-2003
DEFINITION  Adeno-associated virus serum type 1 nucleic acid sequence,
            and host cell containing the same.
ACCESSION   BD242771
VERSION     BD242771.1  GI:33052541
KEYWORDS    JP 2002529098-A/6.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2211)
  AUTHORS   Wilson,J.M. and Xiao,W.
  TITLE     Adeno-associated virus serum type 1 nucleic acid sequence, vector
            and host cell containing the same
  JOURNAL   Patent: JP 2002529098-A 6 10-SEP-2002;
            THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT     OS  AAV-1
            PN  JP 2002529098-A/6
            PD  10-SEP-2002
            PF  02-NOV-1999 JP 2002529098-A/6
            PR  02-NOV-1999 JP          60/107114
            PI  05-NOV-1998 US          60/107114
            PI  JAMES M WILSON,WEIDONG XIAO
            PC  C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
            PC  C12N15/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
            PC  C12N1/21,
            Adeno-associated virus serum type 1 nucleic acid sequence,       CC
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        December 11, 2005, 23:51:42 ; Search time 653 Seconds
                                              (without alignments)
                                              2701.143 Million cell updates/sec

Title:         US-10-696-282-1
Perfect score: 4718
Sequence:      1 ttgcccactccctctctgcg..........cgcagaggaggagtgggcaa 4718

Scoring table: IDENTITY NUC
               Gapop 10.0, Gapext 1.0

Searched:      3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters:       6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Published_Applications_NA_New:*
               1:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
               2:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
               3:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
               4:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
               5:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
               6:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
               7:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
               8:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
               9:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
              10:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3104.6	65.8	4679	7	US-11-184-380-25	Sequence 25, Appl
2	2609.4	55.3	4768	6	US-11-058-751-8	Sequence 1, Appli
3	1444.8	30.6	1872	6	US-10-719-311-1	Sequence 3, Appli
4	1444.8	30.6	1872	6	US-10-719-311-3	Sequence 1, Appl
5	1354.2	28.7	4652	7	US-11-184-380-1	Sequence 15, Appl
6	1266.2	26.8	1611	6	US-10-719-311-14	Sequence 14, Appl
7	935.4	19.8	2208	6	US-10-719-311-15	Sequence 1, Appl
8	883.4	18.7	1197	6	US-10-719-311-13	Sequence 13, Appl
9	734	15.6	2307	7	US-11-184-380-7	Sequence 7, Appli
10	729.8	15.5	2264	7	US-11-184-380-8	Sequence 8, Appli
11	729.8	15.5	2264	7	US-11-184-380-9	Sequence 9, Appli
12	707.8	15.0	939	6	US-10-719-311-12	Sequence 12, Appl
13	707.8	15.0	1690	7	US-11-184-380-15	Sequence 15, Appl
14	690.6	14.6	1870	7	US-11-184-380-11	Sequence 11, Appl
15	630.6	13.4	1800	6	US-10-719-311-17	Sequence 17, Appl
16	588.2	12.5	1617	6	US-10-719-311-19	Sequence 19, Appl
17	476	10.1	1115	7	US-11-184-380-13	Sequence 13, Appl
18	475	10.1	1292	7	US-11-184-380-10	Sequence 10, Appl
19	107.4	2.3	165	7	US-11-058-751-7	Sequence 7, Appli
20	107.4	2.3	272	7	US-11-058-751-6	Sequence 6, Appli
21	105.0	2.2	272	7	US-11-058-751-4	Sequence 4, Appli
22	102.6	2.2	272	7	US-11-058-751-5	Sequence 5, Appli
23	101.8	2.2	4679	7	US-11-184-380-25	Sequence 25, Appl
24	99.4	2.1	165	7	US-11-058-751-7	Sequence 7, Appli
25	97.8	2.1	272	7	US-11-058-751-4	Sequence 4, Appli
26	97.6	2.1	4768	6	US-11-058-751-8	Sequence 1, Appli
27	95.6	2.0	334	7	US-11-058-751-14	Sequence 14, Appl
28	95.6	2.0	334	7	US-11-058-751-14	Sequence 14, Appl
29	95.6	2.0	505	7	US-11-058-751-3	Sequence 3, Appli
30	95.6	2.0	505	7	US-11-058-751-3	Sequence 3, Appli
31	83.4	1.8	125	6	US-10-719-311-6	Sequence 6, Appli
32	83.4	1.8	345	7	US-11-058-751-9	Sequence 9, Appli
33	81.8	1.7	125	6	US-10-719-311-6	Sequence 6, Appli
34	78.6	1.7	345	7	US-11-058-751-9	Sequence 9, Appli
35	78.4	1.7	300	7	US-11-058-751-5	Sequence 5, Appli
36	72	1.5	125	6	US-10-719-311-6	Sequence 6, Appli
37	68	1.4	145	7	US-11-184-380-16	Sequence 16, Appli
38	59	1.3	245	6	US-10-719-311-7	Sequence 7, Appli
39	58	1.2	129	6	US-10-719-311-20	Sequence 20, Appli
40	58	1.2	129	6	US-10-719-311-20	Sequence 20, Appli
41	56.8	1.2	276	7	US-11-058-751-10	Sequence 10, Appli
42	55.2	1.2	276	7	US-11-058-751-10	Sequence 10, Appli
43	55.2	1.2	310	7	US-11-058-751-13	Sequence 13, Appli
44	54.8	1.2	316	7	US-11-058-751-11	Sequence 11, Appli
45	53.6	1.1	310	7	US-11-058-751-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-11-184-380-25

;  Sequence 25, Application US/11184380
;  Publication No. US20050255089A1
;  GENERAL INFORMATION:
;  APPLICANT: Chiorini, John
;  APPLICANT: Kotin, Robert M.
;  TITLE OF INVENTION: AAV5 NUCLEIC ACIDS
;  FILE REFERENCE: 14014.0323U3
;  CURRENT APPLICATION NUMBER: US/11/184,380
;  CURRENT FILING DATE: 2005-07-19
;  PRIOR APPLICATION NUMBER: PCT/US99/11958
;  PRIOR FILING DATE: 1999-05-28
;  PRIOR APPLICATION NUMBER: 60/087,029
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 25
;    LENGTH: 4679
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence; Note =
;    OTHER INFORMATION: synthetic construct
US-11-184-380-25

  Query Match             65.8%;  Score 3104.6;  DB 7;  Length 4679;
  Best Local Similarity   80.2%;  Pred. No. 0;  Matches 884;  Indels 53;  Gaps 10;
  Matches 3788;  Conservative   0;  Mismatches   0;

Qy        1 TTGCCCACTCCCTCTCTGCGCGCGCCTCGCTCGGTGTGGGGCTGCGGACCAAAAGGTCCGC 60
            ||| ||||||||||||||||| |||||  ||||| |||||||| ||| ||||||||||||
Db        1 TTGGCCACTCCCTCTTCTGCGCGCGCGCCTCGGCTCGGTGTGGGGCTGCGGACCAAAGGTCGGCC 60

Qy       61 AGACGGCAGAGCTCTGCTCTTGCCGCGCCCCACCGACGCGAGCGGCAGAGAGGCGGAGTG 120
            ||| ||||||||||||  |||| ||| ||||||  ||||||||||||||||||||||||
Db       61 CGACGCCCGGCGCTTTGCCCGGCGCGCCTCAGTGACGCGAGCCCGGCAGAGAGGCGGAGTG 120

Qy      121 GGCAACTCCATCACTAGGGGTGGGGTAAATGCGGAAGCGCGCTCCCACGCTCACGCGCTGA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db      121 GCCAACTCCATCACTAGGGGGT-----------------------TCCTGGAGGGGTGGGGT 163

Qy      181 CGTAAAATTAGTTCATTAGGG---GAGTGGTTCCTGTATTAGCTGTTCACGTGAGTGCTTTTGC 237
            |||| |||||||||||||||||        ||||| |||||||||| |||||||||||||||||
Db      164 CGTGAATTAGTTCATTAGGGGTTTAGGGAGGGTTCCTGTATTAGAGGTTCACGTGAGTG-TTTTGC 222
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: December 11, 2005, 19:21:31 ; Search time 3212 Seconds
(without alignments)
12146.614 Million cell updates/sec

Title: US-10-696-282-1
Perfect score: 4718
Sequence: 1 ttgccaactcccctctctgcg.........cgcagagaggagtgggcaa 4718

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413469005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   Published Applications NA Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4718	100.0	4718	6	US-10-291-583-6	Sequence 6, Appli
2	4718	100.0	4718	7	US-10-696-261-1	Sequence 1, Appli
3	4718	100.0	4718	7	US-10-696-282-1	Sequence 1, Appli
4	4718	100.0	4718	7	US-10-696-900-1	Sequence 1, Appli
5	4718	100.0	4718	9	US-10-427-129-1	Sequence 1, Appli
6	4718	100.0	4718	9	US-10-959-017-3	Sequence 3, Appli
7	4258	90.3	4683	7	US-10-427-129-6	Sequence 6, Appli
8	4258	90.3	4683	9	US-10-959-017-2	Sequence 2, Appli
9	4253.2	90.1	4683	7	US-10-696-261-19	Sequence 19, Appl
10	4253.2	90.1	4683	7	US-10-696-282-19	Sequence 19, Appl
11	4253.2	90.1	4683	6	US-10-696-900-19	Sequence 19, Appl
12	3947.4	83.7	4721	6	US-10-291-583-11	Sequence 11, Appli
13	3732.4	79.1	13804	7	US-10-427-129-11	Sequence 11, Appli
14	3621	76.7	4382	8	US-10-496-799-1	Sequence 1, Appl
15	3587.2	76.0	4385	6	US-10-291-583-5	Sequence 5, Appl
16	3546.4	75.2	4393	6	US-10-291-583-4	Sequence 4, Appli
17	3546.4	75.2	4393	6	US-10-423-704A-1	Sequence 1, Appli
18	3211.6	68.1	4726	9	US-10-291-583-8	Sequence 8, Appli
19	3211.6	68.1	4726	9	US-10-959-017-4	Sequence 4, Appli
20	3205	67.9	4722	7	US-10-427-129-3	Sequence 3, Appli
21	3107.8	65.9	4679	9	US-09-804-898-1	Sequence 1, Appli
22	3107.8	65.9	4679	3	US-09-945-681-10	Sequence 10, Appl
23	3107.8	65.9	4679	5	US-10-038-972A-12	Sequence 12, Appl
24	3107.8	65.9	4679	6	US-10-136-819-6	Sequence 6, Appli
25	3107.8	65.9	8698	3	US-09-770-315-2	Sequence 2, Appli
26	3084.2	65.4	4681	6	US-10-696-261-18	Sequence 18, Appl
27	3084.2	65.4	4681	7	US-10-696-282-18	Sequence 18, Appl
28	3084.2	65.4	4681	7	US-10-696-900-18	Sequence 18, Appl
29	3082.4	65.3	4680	7	US-10-077-294-1	Sequence 1, Appli
30	3082.4	65.3	4680	5	US-10-163-886-1	Sequence 1, Appli
31	3082.4	65.3	4680	5	US-10-263-127-1	Sequence 1, Appli
32	3082.4	65.3	4680	5	US-10-375-777-1	Sequence 1, Appli
33	3082.4	65.3	4680	10	US-11-063-903-1	Sequence 1, Appli
34	3055.8	64.8	4675	3	US-09-782-378A-1	Sequence 1, Appli
35	3055.8	64.8	4675	3	US-09-782-378A-2	Sequence 2, Appli
36	3055.8	64.8	4675	5	US-10-240-198-1	Sequence 1, Appli
37	3055.8	64.8	4675	6	US-10-291-583-7	Sequence 7, Appli
38	3055.8	64.8	4675	7	US-10-427-129-2	Sequence 2, Appli
39	2993.4	63.4	7557	3	US-10-770-315-3	Sequence 3, Appli
40	2941.8	62.4	8179	5	US-10-205-942-5	Sequence 5, Appli
41	2898.6	61.4	12136	5	US-10-427-129-15	Sequence 15, Appl
42	2730	57.9	8151	5	US-10-205-942-2	Sequence 2, Appli
43	2609.4	55.3	4768	3	US-10-719-311-1	Sequence 1, Appli
44	2609.4	55.3	4768	3	US-10-719-311-1	Sequence 1, Appl
45	2598.6	55.1	4767	7	US-10-427-129-4	Sequence 4, Appli

RESULT 1
US-10-291-583-6
Sequence 6, Application US/10291583
Publication No. US20030138772A1
GENERAL INFORMATION:
APPLICANT: Gao, Guangping
APPLICANT: Wilson, James M.
APPLICANT: Alvira, Mauricio
TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Vir
TITLE OF INVENTION: Sequences and Isolating No. US2003013877A2A1el Sequences Iden
FILE REFERENCE: UPN-0273USA
CURRENT APPLICATION NUMBER: US/10/291,583
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US 60/350,607
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/341,117
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: US 60/377,066
PRIOR FILING DATE: 2002-05-01
PRIOR APPLICATION NUMBER: US 60/386,675
PRIOR FILING DATE: 2002-06-05
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 4718
TYPE: DNA
ORGANISM: adeno-associated virus serotype 1
US-10-291-583-6

Query Match 100.0%; Score 4718; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4718; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

```
Qy    1 TTGCCACTCCCCTCTCTGCGGCGCTCGCTCGCTCGCTCGGTGGGGGGGACCAAAGGTCCGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TTGCCACTCCCCTCTCTGCGGCGCTCGCTCGCTCGCTCGGTGGGGGGGACCAAAGGTCCGC 60

Qy   61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGCGCGCGAGAGAGGGAGTG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGCGCGCGAGAGAGGGAGTG 120

Qy  121 GGCAAACTCCATCACTAGGGGTAATCGCAAAGCGCCTCCCACGCTCGCGCGGTCAGCGGTGA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GGCAACTCCATCACTAGGGGTAATCGCCGAAGCGCCTCCCACGCTCGCGCGGTCAGCGGTGA 180

Qy  181 CGTAAATTACGTCATTAGGGGGAGTGGGGTCCTCGTGTATTAGCTGTCAGTGAGTGCTTTTGCGAC 240
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 12:18:11 ; Search time 810 Seconds
(without alignments)
10353.748 Million cell updates/sec

Title: US-10-696-282-1
Perfect score: 4718
Sequence: 1 ttcgcccactccctctgcg..........cgcgaggaggagtcgggcaa 4718

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1: /cgn2_6/ptcdata/1/ina/1_COMB.seq:*
           2: /cgn2_6/ptcdata/1/ina/5_COMB.seq:*
           3: /cgn2_6/ptcdata/1/ina/6A_COMB.seq:*
           4: /cgn2_6/ptcdata/1/ina/6B_COMB.seq:*
           5: /cgn2_6/ptcdata/1/ina/H_COMB.seq:*
           6: /cgn2_6/ptcdata/1/ina/PCTUS_COMB.seq:*
           7: /cgn2_6/ptcdata/1/ina/PP_COMB.seq:*
           8: /cgn2_6/ptcdata/1/ina/RE_COMB.seq:*
           9: /cgn2_6/ptcdata/1/ina/backfile1.seq:*
              Issued_Patents_NA:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4718	100.0	4718	3	US-09-807-802A-1	Sequence 1, Appli
2	4253.2	90.1	4683	3	US-09-807-802A-19	Sequence 19, Appl
3	3107.8	65.9	4679	3	US-10-038-972A-12	Sequence 12, Appl
4	3107.8	65.9	8698	3	US-09-770-315-2	Sequence 2, Appli
5	3084.2	65.4	4681	3	US-09-807-802A-18	Sequence 18, Appl
6	3082.4	65.3	4680	2	US-08-254-358-1	Sequence 1, Appli
7	3082.4	65.3	4680	2	US-08-475-391-1	Sequence 1, Appli
8	3082.4	65.3	4680	6	PCT-US95-07178-1	Sequence 1, Appli
9	3082.4	65.3	4680	2	US-08-709-609-1	Sequence 1, Appli
10	3055.8	64.8	4675	3	US-09-782-378A-1	Sequence 2, Appli
11	3055.8	64.8	4675	3	US-09-782-378A-2	Sequence 2, Appli
12	2993.4	63.4	7557	3	US-09-770-315-3	Sequence 3, Appli
13	2941.8	62.4	8179	3	US-09-438-268-5	Sequence 5, Appli
14	2730	57.9	8151	3	US-09-438-268-2	Sequence 2, Appli
15	2596	55.0	4767	3	US-09-532-594B-1	Sequence 1, Appli
16	2446.2	51.8	7214	3	US-09-438-268-1	Sequence 1, Appli
17	2211	46.9	2211	3	US-09-807-802A-12	Sequence 12, Appl
18	2204.6	46.7	7447	3	US-10-216-870-11	Sequence 11, Appl
19	1872	39.7	1872	3	US-09-807-802A-4	Sequence 4, Appli
20	1800	38.2	1800	3	US-09-807-802A-14	Sequence 14, Appl
21	1691.2	35.8	4072	3	US-09-770-315-4	Sequence 4, Appli
22	1605	34.0	1605	3	US-09-807-802A-16	Sequence 16, Appl
23	1592	33.7	1641	3	US-09-807-802A-6	Sequence 6, Appli
24	1444.8	30.6	1872	3	US-09-532-594B-3	Sequence 3, Appli
25	1444.8	30.6	1872	3	US-09-532-594B-15	Sequence 15, Appl
26	1425.8	30.2	1872	3	US-08-331-384-2	Sequence 2, Appli
27	1425.8	30.2	4910	2	US-08-836-087-2	Sequence 2, Appli
28	1425.8	30.2	4910	2	US-09-246-320-2	Sequence 2, Appli
29	1425.8	30.2	4910	2	US-09-546-738-2	Sequence 2, Appli
30	1354.2	28.7	4652	3	US-09-533-427-1	Sequence 1, Appli
31	1266.2	26.8	1611	3	US-09-532-594B-14	Sequence 14, Appl
32	1200	25.4	1200	3	US-09-807-802A-8	Sequence 8, Appli
33	1057	22.4	7744	3	US-10-216-870-14	Sequence 14, Appl
34	935.4	19.8	2208	3	US-09-533-594B-5	Sequence 5, Appli
35	920	19.5	969	3	US-09-807-802A-10	Sequence 10, Appl
36	883.4	18.7	1197	3	US-09-532-594B-13	Sequence 13, Appl
37	734	15.6	2307	3	US-09-533-427-7	Sequence 7, Appli
38	729.8	15.5	2264	3	US-09-533-427-3	Sequence 3, Appli
39	729.8	15.5	2264	3	US-09-533-427-9	Sequence 9, Appli
40	707.8	15.0	939	3	US-09-533-594B-12	Sequence 12, Appl
41	690.6	14.6	1690	3	US-09-533-427-15	Sequence 15, Appl
42	690.6	14.6	1870	3	US-09-533-427-11	Sequence 11, Appl
43	630.6	13.4	1800	3	US-09-532-594B-17	Sequence 17, Appl
44	588.2	12.5	1617	3	US-09-532-594B-19	Sequence 19, Appl
45	476	10.1	1115	3	US-09-533-427-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-807-802A-1

Sequence 1, Application US/09807802A
Patent No. 6759237
GENERAL INFORMATION:
APPLICANT: Wilson, James M.
APPLICANT: Xiao, Weidong
TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
TITLE OF INVENTION: Vectors and Host Cells Containing Same
FILE REFERENCE: GNVPN.031USA
CURRENT APPLICATION NUMBER: US/09/807,802A
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/107,114
PRIOR FILING DATE: 1998-11-05
PRIOR APPLICATION NUMBER: PCT/US99/25694
PRIOR FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
LENGTH: 4718
TYPE: DNA
ORGANISM: AAV-1
FEATURE:
NAME/KEY: CDS
LOCATION: (335)..(2206)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (2223)..(4430)
OTHER INFORMATION:
US-09-807-802A-1

Query Match 100.0%; Score 4718; DB 3; Length 4718;
Best Local Similarity 100.0%; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Matches 4718;

```
Qy    1 TTGCCCACTCCCTCTGCGCCCCCGCGCCCTCCGCTCGCTCGGTGGGGGGCTGCGGACCAAAAGGTCCGC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TTGCCCACTCCCTCTGCGCCCCCGCGCCCTCCGCTCGCTCGGTGGGGGGCTGCGGACCAAAAGGTCCGC 60

Qy   61 AGACGGCAGAGACTCGGCTCTGTCCGCGCCCCCACCGAGCGGAGCGGCGCGCAGAGAGGGAGAGTG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AGACGGCAGAGACTCGGCTCTGTCCGCGCCCCCACCGAGCGGAGCGGCGCGCAGAGAGGGAGAGTG 120

Qy  121 GGCAACTCCATCACTAGGGGGTTAATCGCGAAGCGCCTCCCACGGCTGCGGCGGTCAGCGCCTGA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GGCAACTCCATCACTAGGGGGTTAATCGCGAAGCGCCTCCCACGGCTGCGGCGGTCAGCGCCTGA 180
```

OM nucleic - nucleic search, using sw model

```
                                        GenCore version 5.1.6
                                        Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:         December 11, 2005, 12:04:41 ; Search time 16217 Seconds
                                        (without alignments)
                                        13611.732 Million cell updates/sec

Title:          US-10-696-282-1
Perfect score:  4718
Sequence:       1 ttgcccactcctctctgcg...........cgcagagaggagtgggcaa 4718
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       41078325 seqs, 2339541228 residues
Total number of hits satisfying chosen parameters:    82156650
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_est3:*
                 4: gb_htc:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_est7:*
                 9: gb_gss1:*
                10: gb_gss2:*
                11: gb_gss3:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result                  %
No.    Score   Query    Match DB  ID          Description
               Length
-----  -----   -----    -----     --          -----------
  1    180.8   3.8      264   7   CO892248     CO892248  BovGen_20
  2    126.8   2.7      195   7   CO888893     CO888893  BovGen_17
  3     91.8   1.9      489  10   AZ289042     AZ289042  RPCI-23-5
  4     85.4   1.8      778  10   AG491391     AG491391  Mus muscu
  5     76.2   1.6      753   9   BH115587     BH115587  RPCI-24-3
  6     71      1.5      495   6   A2028362     A2028362  RPCI-23-3
  7     60.6    1.3      668   7   CF843972     CF843972  pSHB044xC
  8     55.6    1.2      636   5   BM440166     BM440166  pgrln.pk0
  9     55.2    1.2      659   5   BM371945     BM371945  EBma08_SQ
 10     55.2    1.2      592   5   BQ766288     BQ766288  EBrco8_SQ
 11     55.2    1.2      852   2   BG310300     BG310300  HVSMEc001
 12     55.2    1.2      853   8   BP170704     BP170704  BP170704
 13     55.2    1.2      1553  8   DN551925     DN551925  pSPR-RT-M
 14     54.6    1.2      845   7   CK462848     CK462848  933671_MA
 15     54.4    1.1      665   5   BW846578     BW846578  BW846578
 16     54.4    1.1      638   3   CA024206     CA024206  HZ48Kl3r
 17     54      1.1      415   5   BJ550744     BJ550744  BJ550744
 18     54      1.1      652   3   BJ268130     BJ268130  BJ268130
 19     54      1.1      734   5   BU099382     BU099382  WHE3306_G
 20     54      1.1      818   9   B2578035     B2578035  msh2_5689
 21     54      1.1      1125  8   DR736937     DR736937  FGAS08230
 22     53.8    1.1      878   6   CD437318     CD437318  BL01N0371
```

ALIGNMENTS

```
RESULT 1
CO892248/c
LOCUS       CO892248               264 bp    mRNA    linear   EST 01-SEP-2004
DEFINITION  BovGen 20573 normal cattle brain Bos taurus cDNA clone
            RZPDp1056M03060Q 5', mRNA sequence.
ACCESSION   CO892248
VERSION     CO892248.1  GI:51822548
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 264)
  AUTHORS   Hennig,S., Janitz,M., Herwig,R. and Williams,J.
  TITLE     Generation, annotation, evolutionary analysis and database
            integration of 14969 cattle EST clusters
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Hennig S
            laboratory 123, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1612
            Fax: +49 30 8413 1380
            Email: hennig@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones that display the same hybridisation matrix with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per ONFP cluster was selected for sequencing. cDNA clones and
            filters are distributed via Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH (http://www.rzpd.de).
            PCR Primers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3'
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP) 3'-seq
            Seq primer: 5'-CCGGTCCGGAATTCCGGGT-3' (M13RSP).
FEATURES             Location/Qualifiers
     source          1..264
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /clone="RZPDp1056M03060Q"
                     /sex="female"
                     /tissue_type="brain tissue"
                     /dev_stage="adult brain"
```

```
23    53.2   1.1      556   10   CW491633     CW491633  fsbb001f2
24    53.2   1.1      607    5   BQ660503     BQ660503  H10200Su
25    53.2   1.1      700   10   CW041764     CW041764  104 275 1
26    53     1.1      714    7   CF846078     CF846078  psHB036xD
27    52.6   1.1      484    9   CE239279     CE239279  tigr-gss-
28    52.6   1.1      694    7   CF843224     CF843224  psHB024xE
29    52.6   1.1      818    6   CB905543     CB905543  tric022xj
30    52.6   1.1      528    7   CF870364     CF870364  tric022xj
31    52.4   1.1      818    8   AL818251     AL818251  tric022xj
32    52.4   1.1      688    3   BM816637     BM816637  HB02E06_T
33    52.4   1.1      837    8   CV776978     CV776978  FGAS07138
34    52.4   1.1     1065    4   AY110092     AY110092  Zea mays
35    52.4   1.1     1092    8   DR741775     DR741775  FGAS03081
36    52.2   1.1      838    7   CK371427     CK371427  zmrww005
37    52.2   1.1      942   10   CG329376     CG329376  OGWLJ43TV
38    52     1.1      729   10   CL971124     CL971124  OsIFCC021
39    52     1.1      859    6   CA174679     CA174679  SCJFST101
40    52     1.1      877    3   BI952167     BI952167  HVSME0000
41    51.8   1.1      645    6   CB479173     CB479173  Jn846_G11
42    51.8   1.1      492    6   CF487798     CF487798  PCL1_46_A
43    51.6   1.1      570    2   BB452865     BB452865  89406GD03
44    51.6   1.1      577    5   CF487880     CF487880  PCL1_46_A
45    51.6   1.1      612    3   BI873856     BI873856  9631lIF11
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2005, 10:59:26 ; Search time 2451 Seconds
(without alignments)
12829.057 Million cell updates/sec

Title: US-10-696-282-1
Perfect score: 4718
Sequence: 1 tcgcccactcctctcgcg..........cgcgagagaggagtcgggcaa 4718

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Database : N_Geneseq_21:*
  1: geneseqn1980s:*
  2: geneseqn1990s:*
  3: geneseqn2000s:*
  4: geneseqn2001as:*
  5: geneseqn2001bs:*
  6: geneseqn2002as:*
  7: geneseqn2002bs:*
  8: geneseqn2003as:*
  9: geneseqn2003bs:*
 10: geneseqn2003cs:*
 11: geneseqn2003ds:*
 12: geneseqn2004as:*
 13: geneseqn2004bs:*
 14: geneseqn2005s:*

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	4718	100.0	4718	3	AAD00772	Aad00772 Adeno-ass
2	4718	100.0	4718	10	ADE76507	Ade76507 Adeno-ass
3	4718	100.0	4718	10	ADL13984	Adl13984 Adeno-ass
4	4718	100.0	4718	12	ADG39758	Adg39758 AAV-1 gen
5	4258	90.3	4683	4	AAF23749	Aaf23749 AAv6 DNA
6	4258	90.3	4683	10	ADL13983	Adl13983 Adeno-ass
7	4258	90.3	4683	12	ADG39763	Adg39763 AAV-6 gen
8	3947.4	83.7	4721	10	ADE76502	Ade76502 Adeno-ass
9	3947.4	83.7	4721	12	ADG39764	Adg39764 AAV-7 gen
10	3947.4	83.7	4721	14	ADV67509	Adv67509 Nucleotid
11	3947.4	83.7	4721	14	ADZ27030	Adz27030 Adeno-ass
12	3890	82.5	4347	13	ADW39398	Adw39398 Adeno-ass
13	3890	82.5	4347	14	ADW39398	Adw39398 Adeno-ass
14	3666.8	77.7	4239	13	ADZ46594	Adz46594 HSV-AAV 8
15	3666.8	77.7	4239	14	ADZ46598	Adz46598 HSV-AAV 8
16	3622.6	76.8	4239	14	ADW39402	Adw39402 Adeno-ass
17	3587.2	76.0	4385	10	ABQ80410	Abq80410 AAV9 rep
18	3546.4	75.2	4393	9	ACF05811	Acf05811 Adeno ass
19	3546.4	75.2	4393	10	ADE76505	Ade76505 Adeno-ass
20	3546.4	75.2	4393	12	ADG39765	Adg39765 AAV-8 gen
21	3546.4	75.2	4393	14	ADV67510	Adv67510 Nucleotid
22	3546.4	75.2	4393	14	ADZ27032	Adz27032 Adeno-ass
23	3211.6	68.1	4726	10	ADE76509	Ade76509 Adeno-ass
24	3211.6	68.1	4726	10	ADL13985	Adl13985 Adeno-ass
25	3211.6	68.1	4726	12	ADG39759	Adg39759 AAV-3A ge
26	3205	67.9	4722	4	AAF23748	Aaf23748 AAV3B DNA
27	3205	67.9	4722	12	ADG39760	Adg39760 AAV-3B ge
28	3107.8	65.9	4679	4	AAI66974	Aai66974 Adeno-ass
29	3107.8	65.9	4679	6	ABK89694	Abk89694 Adeno-ass
30	3107.8	65.9	4679	10	ABV76133	Abv76133 Adeno-ass
31	3107.8	65.9	8698	4	AEA48850	Aea48850 Bovine ad
32	3106.2	65.8	8698	4	AAH26325	Aah26325 Wild-type
33	3096.6	65.6	4679	4	AAF23750	Aaf23750 AAV2 DNA
34	3082.4	65.3	4679	12	ADG39757	Adg39757 AAV-2 gen
35	3082.4	65.3	4680	2	AAT09008	Aat09008 Wild-type
36	3082.4	65.3	4680	4	ABX14497	Abx14497 Wild-type
37	3082.4	65.3	4680	9	ADA50070	Ada50070 Adeno-ass
38	3082.4	65.3	4680	10	ADE81075	Ade81075 Adeno-ass
39	3082.4	65.3	4680	12	ADG92079	Adg92079 Wild-type
40	3082.4	65.3	4680	14	AEB48083	Aeb48083 Adeno-ass
41	3055.8	65.3	4675	4	AAH41481	Aah41481 Adeno-ass
42	3055.8	64.8	4675	4	AAF89931	Aaf89931 Nucleotid
43	3055.8	64.8	4675	6	ABA02989	Aba02989 Adeno-ass
44	3055.8	64.8	4675	6	ABS69879	Abs69879 Human ade
45	3055.8	64.8	4675	6	ABS69880	Abs69880 Human ade

ALIGNMENTS

RESULT 1
AAD00772

```
ID   AAD00772 standard; DNA; 4718 BP.
XX
AC   AAD00772;
XX
DT   08-SEP-2000 (first entry)
XX
DE   Adeno-associated virus serotype 1 DNA.
XX
KW   Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW   cap protein; recombinant viral vector; gene delivery; gene therapy;
KW   vaccine; transgene; ss.
XX
OS   Adeno associated virus serotype 1.
XX
FH   Key             Location/Qualifiers
FH
FT   repeat_unit     1. .143
FT                   /label= a
FT                   /note= "inverted terminal repeat which is capable of
FT                   forming T-shaped hairpin structure"
FT   repeat_unit     89. .110
FT                   /*tag= b
FT   protein_bind    124. .125
FT                   /*tag= c
FT                   /bound_moiety= "Rep protein"
FT   misc_feature    219. .226
FT                   /*tag= c
FT                   /note= "Terminal resolute site (TRS)"
FT   misc_binding    219. .226
FT                   /*tag= d
FT                   /bound_moiety= d
FT   promoter        236. .299
FT                   /note= "B box"
FT                   /*tag= e
FT   protein_bind    237. .245
FT                   /*tag= e
FT                   /label= P5_promoter
FT                   /note= "USF"
FT   protein_bind    270. .275
FT                   /*tag= f
FT                   /bound_moiety= "YY1 factor"
FT   TATA_signal     299. .306
FT                   /*tag= g
FT                   /label= P5 TATA-Box
FT   misc_feature
```

QM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run On: December 11, 2005, 12:02:06 ; Search time 22657 Seconds
                           (without alignments)
                           11836.845 Million cell updates/sec

Title: US-10-696-282-1
Perfect score: 4718
Sequence: 1 ttgcccactccctctcgcg..........cgcagagaggagtgggcaa 4718

Scoring table: IDENTITY_NUC
                           Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 2842172S653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
                           1: gb_ba:*
                           2: gb_in:*
                           3: gb_env:*
                           4: gb_om:*
                           5: gb_ov:*
                           6: gb_pat:*
                           7: gb_ph:*
                           8: gb_pr:*
                           9: gb_ro:*
                          10: gb_sts:*
                          11: gb_sy:*
                          12: gb_un:*
                          13: gb_vi:*
                          14: gb_htg:*
                          15: gb_pl:*
                              GenEmbl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4718	100.0	4718	6	BD242766	BD242766 Adeno-ass
2	4718	100.0	4718	6	AR562498	AR562498 Sequence
3	4718	100.0	4718	6	AX753251	AX753251 Sequence
4	4718	100.0	4718	13	AF063497	AF063497 Adeno-
5	4258	90.3	4683	13	AF028704	AF028704 Adeno-
6	4253.2	90.1	4683	6	BD242775	BD242775 Adeno-ass
7	4253.2	90.1	4683	6	AR562507	AR562507 Sequence
8	3947.4	83.7	4721	6	CQ972062	CQ972062 Sequence
9	3947.4	83.7	4721	6	CS073592	CS073592 Sequence
10	3947.4	83.7	4721	6	AX753246	AX753246 Sequence
11	3947.4	83.7	4721	13	AF513851	AF513851 Adeno-ass
12	3587.2	76.0	4385	6	AX753250	AX753250 Sequence
13	3546.4	75.2	4393	6	CQ972063	CQ972063 Sequence
14	3546.4	75.2	4393	6	CS073594	CS073594 Sequence
15	3546.4	75.2	4393	6	AX753249	AX753249 Sequence
16	3546.4	75.2	4393	13	AF513852	AF513852 Adeno-
17	3293.8	69.8	4102	13	AY631965	AY631965 Adeno-ass
18	3211.6	68.1	4726	6	AX753253	AX753253 Sequence
19	3211.6	68.1	4726	13	AVU48704	U48704 Adeno-assoc
20	3205	67.9	4722	13	AF028705	AF028705 Adeno-ass
21	3107.8	65.9	4679	13	AX282480	AX282480 Sequence
22	3107.8	65.9	4679	13	AF043303	AF043303 Adeno-
23	3107.8	65.9	8698	6	AR222044	AR222044 Sequence
24	3107.8	65.9	8698	6	AX205072	AX205072 Sequence
25	3084.2	65.4	4681	6	BD242774	BD242774 Adeno-ass
26	3084.2	65.4	4681	6	AR562506	AR562506 Sequence
27	3082.4	65.3	4680	6	AR028767	AR028767 Sequence
28	3082.4	65.3	4680	6	I62303	I62303 Sequence 1
29	3055.8	64.8	4675	6	BD094552	BD094552 Method of
30	3055.8	64.8	4675	6	AX135805	AX135805 Sequence
31	3055.8	64.8	4675	6	AX286292	AX286292 Sequence
32	3055.8	64.8	4675	6	AX753252	AX753252 Sequence
33	3055.8	64.8	4675	13	AA2CG	AA2CG
34	2993.4	63.4	7557	6	AR222045	AR222045 Sequence
35	2993.4	63.4	7557	6	AX205073	AX205073 Sequence
36	2972.6	63.0	4429	13	AY695371	AY695371 Adeno-ass
37	2971	63.0	4429	13	AY695372	AY695372 Adeno-ass
38	2969.6	62.9	4429	13	AY695376	AY695376 Adeno-ass
39	2963.2	62.8	4430	13	AY695374	AY695374 Adeno-ass
40	2961.4	62.8	4429	13	AY695373	AY695373 Adeno-ass
41	2959.8	62.7	4429	13	AY695375	AY695375 Adeno-ass
42	2941.8	62.4	8179	6	BD271148	BD271148 Virus vec
43	2941.8	62.4	8179	6	AR264580	AR264580 Sequence
44	2878.4	61.0	7327	11	AF369963	AF369963 Cloning v
45	2731	57.9	4087	13	AY631966	AY631966 Adeno-ass

ALIGNMENTS

RESULT 1
BD242766
LOCUS   BD242766   4718 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   Adeno-associated virus serum type 1 nucleic acid sequence, vector
               and host cell containing the same.
ACCESSION   BD242766
VERSION   BD242766.1 GI:33052536
KEYWORDS   JP 200529098-A/1.
SOURCE   unidentified
  ORGANISM   unidentified
               unclassified.
REFERENCE   1 (bases 1 to 4718)
  AUTHORS   Wilson,J.M. and Xiao,W.
  TITLE   Adeno-associated virus serum type 1 nucleic acid sequence, vector
               and host cell containing the same
  JOURNAL   Patent: JP 200529098-A 1 10-SEP-2002;
               THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT   AAV-1
PN   JP 200529098-A/1
PD   10-SEP-2002
PF   02-NOV-1999 JP 2000581227
PR   05-NOV-1998 US 60/107114
PI   JAMES M WILSON, WEIDONG XIAO
PC   C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
PC   C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
     Adeno-associated virus serum type 1 nucleic acid sequence, CC
     vector and host
CC   cell containing the same
FH   Key   Location/Qualifiers
FT   source   1..4718
FT   CDS   (335)..(2206)
FT   CDS   (2223)..(4430).

FEATURES   Location/Qualifiers
     source   1..4718
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match   100.0%;   Score 4718;   DB 6;   Length 4718;
Best Local Similarity   100.0%;   Pred. No. 0;